

**Request to Archive
With The National Centers for Environmental Information
For Big Earth Data Initiative (BEDI) Global Monitoring Division (GMD) Aerosol Group:
Long-Term Surface Aerosol Measurements
Provided by NOAA/OAR/ESRL**

2017-05-19

This information will be used by NCEI to conduct an appraisal and make a decision on the request.

1. Who is the primary point of contact for this request?

Rebecca Evrard
GMD
BEDI Coordinator
rebecca.l.evrard@noaa.gov
Working with Heather Brown

2. Name the organization or group responsible for creating the dataset.

NOAA-GMD

3. Provide an overview summarizing the scope of data you want to archive. Describe the outputs, data variables, including their measurement resolution and coverage.

This is to archive 24 data sets under the ESRL/GMD purview. They range from raw to product levels of processing and cover domains such as aerosols; Halocarbons and Other Atmospheric Trace Gases; Ozone and Water Vapor; Global Radiation; and Carbon Cycle and Greenhouse Gases. Each data set will want to receive a DOI.

First dataset to be done is the Aerosol optical properties (scattering, absorption, and number concentration) depending on instrument availability at hourly time resolution. This is time and station dependent data. Some of our stations currently have all of these measurements, while some have just one. Some had more or less of the current set of measurements in the past.

4. What is the time period covered by the dataset? (YYYY-MM-DD, YYYY-MM or YYYY)

From 1974-01-01
Ongoing as continuous updates to the data record

5. Edition or version number(s) of the dataset:

N/A

6. Approximate date when the dataset was or will be released to the public:

2017-06-19

7. Who are the expected users of the archived data? How will the archived data be used?

Science organizations, climate researchers, GCM and CTM modelers looking to reproduce in situ observations.

8. Has the dataset undergone user evaluation and/or an independent review process? Did NCEI participate in design reviews?

Yes. The particle number concentration data have been evaluated and used in a multi-station trends paper by Asmi et al. (2013). The aerosol light scattering data have been evaluated and used in a multi-station trends paper by Collaud Coen et al. (2013).

9. Describe the dataset's relationship to other archived datasets, such as earlier versions or related source data. If this is a new version, how does it improve upon the previous version(s)?

Conversion to netCDF-4 for use with modern tools.

These data are all archived in the World Data Center for Aerosols at NILU in Norway. The only difference is that these data will be converted to a different (NetCDF4) format.

10. List the input datasets and ancillary information used to produce the data.

None

11. List web pages and other links that provide information on the data.

<https://esrl.noaa.gov/gmd/aero/index.html>

Heather Brown's breakdown sheet:

https://docs.google.com/spreadsheets/d/1hR7e-gpbIF970NsSajvSYqRQ_2USXuV7ehjzssquw00/edit#gid=0

12. List the kinds of documents, metadata and code that are available for archiving. For example, data format specifications, user guides, algorithm documentation, metadata compliant with a standard such as ISO 19115, source code, platform/instrument metadata, data/process flow diagrams, etc.

1. Metadata in the netCDF files
2. Source code for conversion of original data files to netCDF

13. Indicate the data file format(s).

1. netCDF-4

14. Are the data files compressed?

No

15. Provide details on how the files are named and how they are organized (e.g., file_name_pattern_YYYYMM.tar in monthly aggregations).

ex: AEROSOL_HOUR_GAWSTATION_STARTISO_ENDISO_GENERATEISO.nc

16. Explain how to access sample data files and/or a file listing for previewing. If it is not available now, when will it be available?

ftp://aftp.cmdl.noaa.gov/aerosol/thd/2015/AEROSOL_HOUR_THD_20150101T000000Z_20160101T000000Z_20161030T051124Z.nc

17. What is the total data volume to be submitted?

Historic Data: all historic data or data submitted as a completed collection.

Total Data Volume: 62GB
Number of Data Files: 1000000

Continuous Data: data volume rate for a continuous data production.

Total Data Volume Rate: 1.5GB per Year
Data File Frequency: 3000 per Year
Data Production Start: 1974-01-01

18. Are later updates, revisions or replacement files anticipated? If so, explain the conditions for submitting these additional data to the archive.

Data would be updated based on station mentor revisions (if any).

This will almost certainly happen for our filter-based light absorption data when a new/better correction algorithm is available.

19. Describe the server that will connect to the ingest server at NCEI for submitting the data.

Physical Location: Boulder, CO

System Name: <http://aero2.cmdl.noaa.gov/>

System Owner: NOAA

Additional Information: Might be pushed from my workstation or above server.

20. What are the possible methods for submitting the data to NCEI? Select all that apply.

1. SFTP PUSH

21. Identify how you would like NCEI to distribute the data. Web access support depends on the resources available for the dataset.

1. User interface to order and stage data for download

2. Direct download links

22. Will there be any distribution, usage, or other restrictions that apply to the data in the archive?

No known constraints apply to the data.

23. Discuss the rationale for archiving the dataset and the anticipated benefits. Mention any risks associated with not archiving the dataset at NCEI.

Make the data available in netCDF-4 format. This data may be found in conjunction with other data sets at NCEI by time, location or other aspects providing value to researchers.

24. Are the data archived at another facility or are there plans to do so? Please explain.

Yes. The data is currently archived at noaa.gov and at WDCA.

25. Is there an existing agreement or requirement driving this request to archive? Have you already contacted someone at NCEI?

BEDI 2016 proposal # 38. We have submitted one file from the collection via Send2NCEI and a file from another collection.

26. Do you have a data management plan for your data?

No

27. Have funds been allocated to archive the data at NCEI?

BEDI 2016 proposal # 38

28. Identify the affiliated research project, its sponsor, and any project/grant ID as applicable.

Aerosols group, NOAA/ESRL/GMD, BEDI 2016 proposal #38

29. Is there a desired deadline for NCEI to archive and provide access to the data?

Archive by: 2017-06-19

Accessible by: 2017-06-19

30. Add any other pertinent information for this request.

We would like to have this collection archived ASAP.